

(19) 日本国特許庁(JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-98655

(43) 公開日 平成10年(1998)4月14日

(51) Int. C1. 6
H 04 N 5/44
G 11 B 15/02 3 2 8
H 04 N 5/445

識別記号
F I
H 04 N 5/44 Z
G 11 B 15/02 3 2 8 S
H 04 N 5/445 Z

審査請求 未請求 請求項の数 3

O L

(全 6 頁)

(21) 出願番号 特願平8-252119
(22) 出願日 平成8年(1996)9月25日

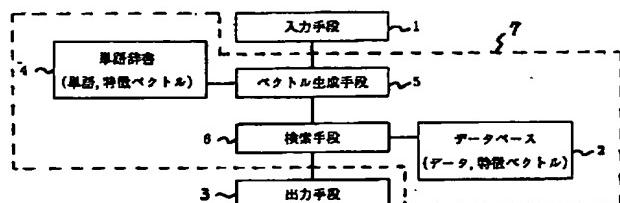
(71) 出願人 000005049
シャープ株式会社
大阪府大阪市阿倍野区長池町22番22号
(72) 発明者 黒武者 健一
大阪府大阪市阿倍野区長池町22番22号 シ
ヤープ株式会社内
(72) 発明者 池内 洋
大阪府大阪市阿倍野区長池町22番22号 シ
ヤープ株式会社内
(72) 発明者 芥子 育雄
大阪府大阪市阿倍野区長池町22番22号 シ
ヤープ株式会社内
(74) 代理人 弁理士 梅田 勝

(54) 【発明の名称】番組検索装置

(57) 【要約】

【課題】 番組検索やビデオ録画予約する際に、手間がかかる、入力ミスが生じたり、検索時間がかっていた。

【解決手段】 番組の日時、チャンネル、番組名、番組内容、出演者等の番組データを記憶したデータベース2を備え、所望の番組に関連する単語を入力手段1から入力して、各単語の特徴ベクトルをもつ単語辞書4を参照して、入力手段1からの質問文に対応する特徴ベクトルをベクトル生成手段5で作成し、その特徴ベクトルと上記データベース中の番組データに対応する特徴ベクトルとの距離を検索手段6で計算して検索する。



【特許請求の範囲】

【請求項1】 所望の番組に関連する単語を入力する入力手段と、
単語と該単語の特徴単語との意味的な結合関係を示す特徴ベクトルとを組にして記憶された単語辞書と、
番組データの単語と該単語の特徴単語との意味的な結合関係を示す特徴ベクトルとを組にして記憶されたデータベースと、
上記入力手段から入力された単語から上記単語辞書を参照して対応する特徴ベクトルを生成するベクトル生成手段と、
該ベクトル生成手段により生成された特徴ベクトルと上記データベースの特徴ベクトルとの距離を計算する検索手段と、
上記距離に基づいて番組データを出力する出力手段とを備えたことを特徴とする番組検索装置。

【請求項2】 上記番組データは、番組の日付、開始時刻、終了時刻、チャンネル、番組名、番組内容等からなり、上記番組データを上記データベースに記憶させるデータ入力手段をさらに備えたことを特徴とする請求項1に記載の番組検索装置。

【請求項3】 上記番組データは、番組の日付、開始時刻、終了時刻、チャンネル、番組名、番組内容等からなり、番組設定や変更、またはビデオ録画予約のために用いられることを特徴とする請求項1に記載の番組検索装置。

【発明の詳細な説明】**【0001】**

【発明の属する技術分野】 本発明は、番組検索装置に関し、テレビ、CATV、文字放送、ラジオ番組などの番組検索やビデオ録画予約に用いられるものである。

【0002】

【従来の技術】 テレビ、CATV、文字放送、ラジオ番組を視聴する場合、またビデオでこれらの番組の録画予約を行う場合、新聞やテレビ情報誌等で番組の時間やチャンネルを調べる必要がある。また、どのような番組が放送されているかも、これら情報誌を調べる必要がある。

【0003】 近年、テレビ放送も地上波放送以外に、BS放送やCS放送が視聴できるようになり、多チャンネル化が進んでおり、CATVや文字放送ではさまざまな番組を視聴することができる。このように、番組数が増えると目的の番組を探すのに手間が掛かり、情報誌を調べるのに時間も掛かることになる。また、これらの放送をビデオに録画予約をする場合、情報誌を見ながら手で入力を行うので、予約内容の入力間違を生ずる場合がある。

【0004】 このような番組を探す従来の技術（従来技術1）として、特開平3-32285号公報に記載されているように、ビデオ番組情報のデータベースを持た

せ、各番組のデータに検索用のキーワードを付けておくことによって、利用者が入力したキーワードを検索してビデオ番組を探すことができるビデオ番組検索装置がある。

【0005】 また、他の番組を探す従来の技術（従来技術2）として、特開平3-88159号公報に記載されているように、番組情報データベースを作成する時に、各番組にスポーツやバラエティ等のジャンルコードを付けておくことによって、番組を検索したい時に、利用者がジャンルコードを指定することによって、目的の予約したい番組を絞ることができる番組予約システムがある。

【0006】 さらに、他の番組を探す従来の技術（従来技術3）として、特開平7-143463号公報に記載されているように、番組を検索する際、入力したキーワードと番組のジャンル名やタイトル名とが一致しなかつた場合でも、ファジー推論用のプログラムを設けることによって、検索を可能にしている文字放送番組検索装置がある。

【0007】

【発明が解決しようとする課題】 しかし、従来技術1では、キーワード検索であるので、入力したキーワードが完全に一致しないければ検索ができないという問題がある。また、従来技術2では、ジャンルコードを設定するだけでは十分に番組を絞り込むことはできず、目的の番組を十分に絞り込むためにはジャンルコードを細かく設定する必要があり、利用者が望む範囲での検索は難しいという問題がある。

【0008】 さらに、従来技術3では、ファジー推論を行るために、番組データベース中の全ての番組について、各キーワードに対するファジー値を計算する必要があり、このファジー値の計算量が莫大になるという問題がある。つまり、番組データベース中の番組数がn、検索キーワード数をmとすると、 $n \times m$ 回のファジー値計算が必要になる。また、新たな番組データをデータベースに付け加える度にm回のファジー値計算が必要になる。番組数を1000、キーワード数を100000とすると、1つの番組に対するファジー値計算回数が10000回、データベース全体に対しては10000000回となる。

【0009】 本発明の目的は、テレビ、CATV、文字放送、ラジオ番組などの番組検索やビデオ録画予約を簡単にを行うことができ、検索キーワードが完全に一致していない場合でも検索を行うことができ、検索のための計算量も小さい、番組検索装置を提供するものである。

【0010】

【課題を解決するための手段】 請求項1に記載の番組検索装置は、所望の番組に関連する単語を入力する入力手段と、単語と該単語の特徴単語との意味的な結合関係を示す特徴ベクトルとを組にして記憶された単語辞書と、

番組データの単語と該単語の特徴単語との意味的な結合関係を示す特徴ベクトルとを組にして記憶されたデータベースと、上記入力手段から入力された単語から上記単語辞書を参照して対応する特徴ベクトルを生成するベクトル生成手段と、該ベクトル生成手段により生成された特徴ベクトルと上記データベースの特徴ベクトルとの距離を計算する検索手段と、上記距離に基づいて番組データを出力する出力手段とを備えたことを特徴とする。

【0011】請求項2に記載の番組検索装置は、請求項1に記載の番組検索装置において、上記番組データは、番組の日付、開始時刻、終了時刻、チャンネル、番組名、番組内容等からなり、上記番組データを上記データベースに記憶させるデータ入力手段をさらに備えたことを特徴とする。

【0012】請求項3に記載の番組検索装置は、請求項1に記載の番組検索装置において、上記番組データは、番組の日付、開始時刻、終了時刻、チャンネル、番組名、番組内容等からなり、番組設定や変更、またはビデオ録画予約のために用いられることを特徴とする。

【0013】

【発明の実施の形態】本発明の番組検索装置は、特徴ベクトルを用いた検索装置を利用している。その特徴ベクトルを用いた番組検索装置の構成を図1に示す。この番組検索装置は、所望の番組の内容に意味的に近い内容の質問文を利用者が入力するキーボードやタブレットやマイクからなる入力手段1と、番組データと各番組データに対応する特徴ベクトルとが組になって記憶されているデータベース2と、検索結果を出力するCRTやプリンタからなる出力手段3と、単語と各単語に対応する特徴ベクトルとが組になって記憶されている単語辞書4と、入力手段1から入力された単語と単語辞書4とからベクトル生成を行うベクトル生成手段5と、ベクトル生成手段5からのベクトル生成結果とデータベース2とから所定の演算を行う検索手段6とから構成される。検索装置7は、汎用のコンピュータやメモリや外部記憶装置で構成される。

【0014】まず、特徴ベクトルについて説明する。特徴ベクトルとは、文章中の単語が持つ概念と文脈との関係を示したものであり、多数の特徴単語との意味的な結合関係の程度をベクトル表現したものである。本発明の特徴ベクトルは、(社)電子情報通信学会発行の信学技法A192-99(1993-1)「大規模データベース」

$$Q \cdot S = 3 \times 4 + 5 \times 5 + 4 \times 4 + 2 \times 1 + 4 \times 4 + 5 \times 5 + 2 \times 0 + 1 \times 1$$

$$= 97$$

$$Q \cdot T = 3 \times 5 + 5 \times 0 + 4 \times 4 + 2 \times 6 + 4 \times 3 + 5 \times 1 + 2 \times 3 + 1 \times 2$$

$$= 68$$

となり、質問文qは、データtよりデータsの方に距離が近いことになる。

【0018】

【実施例】図2に、本発明の実施例に係る番組検索装置

*スからの連想検索」での文脈ベクトルによる文書検索技術を用いる。つまり、本実施の形態での「特徴ベクトル」は上記「文脈ベクトル」にそのまま対応する。

【0015】k個の概念分類を特徴単語とすると、k次元ベクトルの各要素の値を一つ一つの特徴単語に対応させることになる。単語iの文脈ベクトルX_i=(x_{i1}, x_{i2}, ..., x_{ik})の各要素の値は、0≤x_{ij}≤E_mとなる。E_mは、正の定数である。単語iと特徴単語jとの間に関係がない場合には、x_{ij}=0になります。

10 関係がある場合には、その関係の程度に応じて大きい値をとる。例えば、特徴ベクトルが5つの特徴単語(自然、都会、騒音、動物、緑)から成り立っているとし、それぞれの要素の値が0か1の2値である場合には、単語「山」の特徴ベクトルを、

$$(1, 0, 0, 1, 1)$$

などと表すことができる。

【0016】次に、番組検索の手順について説明する。入力手段1から利用者が所望の番組の内容に意味的に近い内容の質問文を入力する。質問文は、単語を含んでいれば、単語ごとに入力しても文章や自然言語で入力しても良い。図7に示すように、ベクトル生成手段5により、入力された質問文が文章等であれば個々の単語が抽出され、単語辞書4からそれぞれの単語に対応する特徴ベクトルを読み出し、それらのベクトルの和を計算して各ベクトルを正規化する。検索手段6では、このベクトルの和とデータベース2中の各番組データに対応する特徴ベクトルとの距離を計算し、最も距離の近い特徴ベクトルを持つ番組データから順に、それらの番組データを検索結果として出力手段3に出力する。

【0017】具体的に、検索手段6での距離計算について説明する。本発明では、距離の計算のために内積計算を行っている。2つのベクトルをX=(x₁, x₂, ..., x_m)、Y=(y₁, y₂, ..., y_m)とすると、内積X·Yは、

$$X \cdot Y = x_1 \times y_1 + x_2 \times y_2 + \dots + x_m \times y_m$$

となる。この内積値が大きいほど距離が近いことになる。例えば、質問文qの特徴ベクトルQが、

$$Q = (3, 5, 4, 2, 4, 5, 2, 1)$$

で、データベース2中のデータs、tの特徴ベクトルS、Tが、

$$S = (4, 5, 4, 1, 4, 5, 0, 1)$$

$$T = (5, 0, 4, 6, 3, 1, 3, 2)$$

とすると、

$$S \cdot Q = 4 \times 3 + 5 \times 5 + 4 \times 4 + 1 \times 2 + 4 \times 4 + 5 \times 5 + 0 \times 1 + 1 \times 1$$

$$= 97$$

$$S \cdot T = 4 \times 5 + 5 \times 0 + 4 \times 4 + 1 \times 6 + 4 \times 4 + 5 \times 3 + 0 \times 1 + 1 \times 2$$

$$= 68$$

を示す。ここで、図1と同一部には同一符号を付す。図1と比較して、検索装置7に、データベース2へのデータを入力するためのデータ入力手段8と、被制御機器1と0を制御したり番組データを送るための有線あるいは無

線の通信手段 9 と、被制御機器 10 とが付加されている。被制御機器 10 は、テレビ、CATV 受像機、文字放送受像機、ラジオ等の受像機あるいは受信機、あるいはビデオである。検索装置 7 は、出力手段 3 に出力された検索結果の番組データの候補から入力手段 1 により選択された番組データの全部または一部を通信手段 9 を介して被制御機器 10 に出力する。

【0019】検索装置 7 には、データ入力手段 8 から各番組の日付、開始時刻、終了時刻、チャンネル、番組名、番組内容、出演者等の番組データが入力され、データベース 2 に記憶されている。この番組データのデータ構成例を図 3 に示す。

【0020】次に、データベース 2 の番組データベースの作成について説明する。データ入力手段 8 からの入力は、キーボードからの入力、OCR 入力、音声入力、パソコン通信等のオンラインデータからの入力、文字放送データ受信等による。入力された番組データは、番組データごとに特徴ベクトルに変換され、番組データと特徴ベクトルとが組になってデータベース 2 に記憶される。ここでの、特徴ベクトルの生成は、ベクトル生成手段 5 の働きにより番組データから個々の単語を抽出し、単語辞書 4 からそれぞれの単語に対応する特徴ベクトルを読み出し、それらのベクトルの和を計算し、ベクトルの大きさが一定になるように正規化することで行う。

【0021】具体的に、この番組データを特徴ベクトルに置き換える処理について説明する。この処理は、単語から特徴ベクトルを生成する処理と同様である。ここで、番組データの内容が図 4、単語辞書 4 の内容の一部が図 5 である場合を考える。ここで、番組データから番組名、番組内容、その他のデータの部分を取り出し、助詞を取り除いて各単語を抽出する。ここでは、「ニュース」、「今日」、「特集」、「春」、「話題」が抽出される。次に、これらの単語の特徴ベクトルを、単語辞書 4 から読み出して、これらのベクトルの和 V_s を計算する。単語 X の特徴ベクトルを $V(X)$ で表すとすると、

$$V_s = V(\text{ニュース}) \times 2 + V(\text{今日}) + V(\text{特集}) + V(\text{春}) + V(\text{話題})$$

となる。「ニュース」は 2箇所で抽出されているので 2 倍されている。図 5 の特徴ベクトルの値を代入してみると、

$$V_s = (3, 5, 0, 4, 2, \dots, 3)$$

となる。次に、このベクトルを正規化する。正規化されたベクトルを V_n とすると、

$$V_n = a \times V_s / |V_s|$$

で計算される。ここで、 a は正規化された後のベクトルの大きさ、 $|V_s|$ はベクトル V_s の大きさで正の値を取り、 $V_s = (v_1, v_2, \dots, v_m)$ とすると、

$$|V_s|^2 = |v_1|^2 + |v_2|^2 + \dots + |v_m|^2$$

である。ここでは、 $a = 10$ として正規化を行つて、このようにして正規化された特徴ベクトルと番組データ

ークとが組になってデータベース 2 の中に記憶される。

【0022】図 6 は、番組検索の処理のフローチャートである。まず、入力手段 1 から検索したい番組内容に意的的に近い番組名や番組内容等の検索条件の質問文を入力する (S1)。検索装置 7 により、データベース 2 中の番組データの検索を行う (S2)。ここで行われる検索は連想検索であるので、番組データ中に完全に一致するデータが存在しなくとも、意的的に近い内容の番組データを検索することができる。検索装置 7 の検索結果を出力手段 3 に出力して (S3)、検索内容の確認および選択を行う (S4)。選択は、例えば CRT に表示された番組データの候補から所望の番組データを入力手段 1 で指定することにより行う。そして、検索結果の中から選択された番組データの内、必要な番組データが通信手段 9 を介して被制御機器 10 に送られる。チャンネル設定や変更の場合 (S5) は、被制御機器 10 に番組データ中のチャンネルのデータが送信され、受け取ったチャンネルに基づいて被制御機器 10 のチャンネルの設定や変更をする。また、番組がまだ始まっている場合に

10 20 は、番組データ中の日付、開始時刻、必要であれば終了時刻のデータを送信して被制御機器 10 に設定する。一方、ビデオ録画予約の場合 (S5) は、被制御機器 10 にビデオ録画予約したい番組データ中のチャンネル、日付、開始時刻、必要であれば終了時刻のデータが送信され、受け取ったデータに基づいて被制御機器 10 のビデオ録画予約の設定をする。ビデオ録画予約時に、番組データを出力手段 3 に出力することによって、番組内容の確認を行うことができるので、ビデオ録画予約内容に間違いがないかを確認することができる。

30 30 【0023】検索装置 7 は特徴ベクトルを用いた検索を行うので、意的的な検索を行って番組を探し出すことが可能である。例えば、「お笑い番組」と入力手段 1 から入力した場合には、番組データの中に「お笑い番組」が含まれているもの以外にも、「コメディ」や「漫才」、「落語」等のデータの含まれた番組も検索することができる。このように、目的の番組の内容を具体的に入力しなくても、意的的に近い内容の番組を検索することができる。さらに、日付、曜日、時刻、出演者、チャンネル等の条件を加えることによって、検索結果を絞ることができ、目的の番組を探し出すことができる。

40 【0024】また、ファジー推論を利用したものと比較して、連想検索の計算量は小さくなっている。データベースの番組データを特徴ベクトルに書き換えるのには、1 つの番組に対しては、その番組データに含まれている単語の特徴ベクトルを単語辞書から読み出して、その番組データに対する特徴ベクトルを計算するだけなので、特徴ベクトルの計算回数は 1 回、データベース全体に対しては、データベース中の番組数を n とすると、 n 回になる。番組数を 1000 とすると、データベース全体に対する特徴ベクトル計算回数は 1000 回となり、単語

7

辞書中の単語数には関係がない。

【0025】

【発明の効果】本発明の番組検索装置によれば、番組検索に特徴ベクトルを用いた意味的な検索を行っているので、目的の番組の内容を具体的に指定しなくても、意味的に近い内容を入力すれば、目的の番組を簡単に検索することができ、番組設定の場合には、番組設定や変更を確実にし、ビデオ録画予約の場合には、予約間違いを防ぐことができる。

【図面の簡単な説明】

【図1】本発明の番組検索装置の構成図である。

【図2】本発明の実施例の構成図である。

【図3】本発明の実施例の番組データの構成図である。

【図4】本発明の実施例の番組データの具体例である。

8

【図5】本発明の実施例の単語辞書の一部を示す図である。

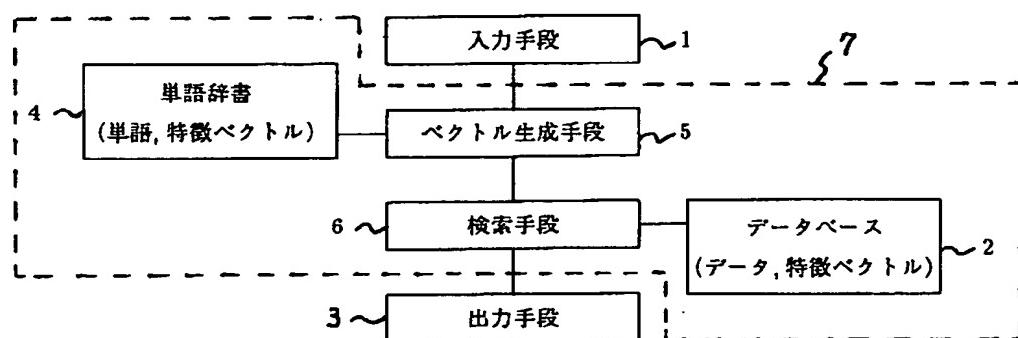
【図6】本発明の実施例の検索のフローチャートである。

【図7】本発明の特徴ベクトル化のフローチャートである。

【符号の説明】

- 1 入力手段
- 2 データベース
- 3 出力手段
- 4 単語辞書
(単語、特徴ベクトル)
- 5 ベクトル生成手段
- 6 検索手段
- 7 データベース
(データ、特徴ベクトル)
- 10 10
- 11 検索手段
- 12 出力手段

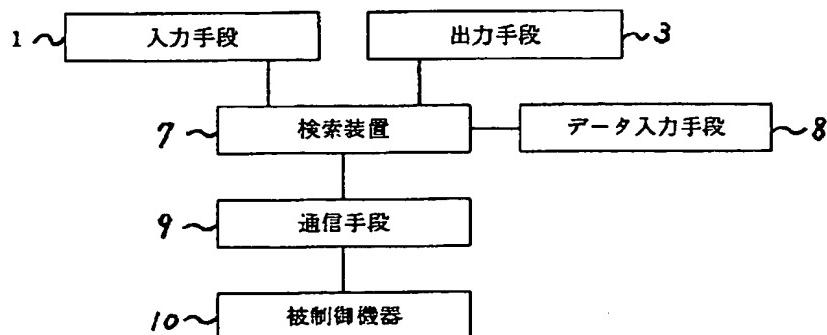
【図1】



【図5】

単語	特徴ベクトル
ニュース	(1, 1, 0, 1, 0, ..., 0)
今日	(1, 0, 0, 1, 1, ..., 0)
特集	(0, 1, 0, 0, 0, ..., 0)
春	(0, 1, 0, 0, 1, ..., 1)
夏	(0, 1, 0, 1, 0, ..., 0)

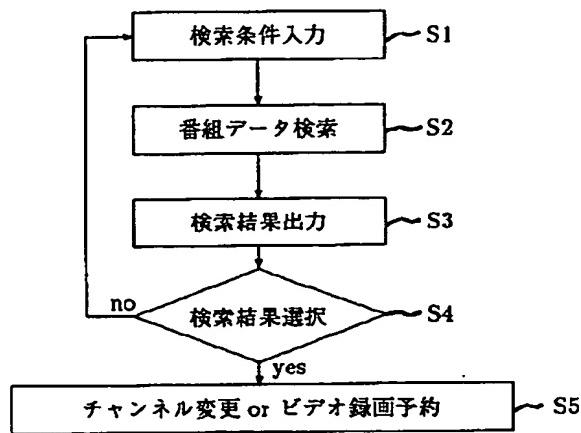
【図2】



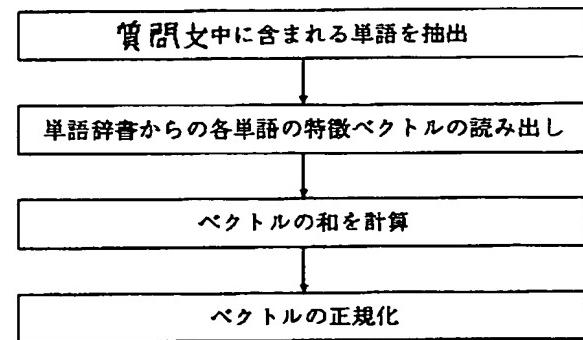
【図3】

【図4】

【図6】



【図7】



PATENT ABSTRACTS OF JAPAN

(11) Publication number : 11-187324
(43) Date of publication of application : 09.07.1999

(51) Int. Cl. HO4N 5/44
HO4H 7/00
HO4N 5/262

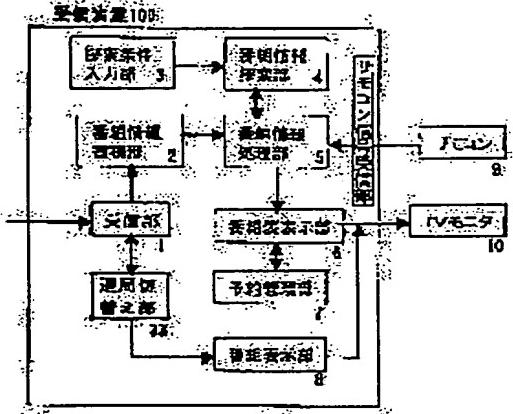
(21) Application number : 09-364346 (71) Applicant : MATSUSHITA ELECTRIC IND CO LTD
(22) Date of filing : 19.12.1997 (72) Inventor : ARAI YUIKO
NAGAO TAKESHI
FUJITA KENICHI
SUZUKI TAKAYUKI

(54) PROGRAM INFORMATION PREPARING DEVICE, ITS METHOD AND RECEIVER

(57) Abstract:

PROBLEM TO BE SOLVED: To provide a receiver capable of displaying a user's own channel.

SOLUTION: This receiver 100 includes a program information storing part 2 storing program information with names, start times and channel identifying information, a program information searching part 4 searching program information according to a holding searing condition and preparing search result program information and a program table display part 6 displaying search result program information as a unique channel. Thereby, the user is own channel is displayed.



LEGAL STATUS

[Date of request for examination] 21.01.2000

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]